JOURNAL OF MULTIVARIATE ANALYSIS 55, 14–28 (1995)

Hadamard Differentiability on $D[0, 1]^p$

JIAN-JIAN REN

University of Nebraska-Lincoln

AND

PRANAB KUMAR SEN

University of North Carolina at Chapel Hill

We show that a statistical functional is asymptotically normal if it induces a Hadamard differentiable functional defined on the space $D[0, 1]^p$. This work involves p-dimensional empirical processes. As an example we illustrate the use of the von Mises method in proving the asymptotic normality of Spearman's rank correlation coefficient. © 1995 Academic Press, Inc.

1. INTRODUCTION

The asymptotic normality of a statistical functional via the Hadamard derivative for univariate observations has been studied by Reeds [7] (1976) and Fernholz [2]. In this paper, we generalize their method to multivariate observations.

If we have a univariate sample $X_1, ..., X_n$ from a distribution F, a functional $T(\cdot)$ of the sample distribution function $F_n(x) = n^{-1} \sum_{i=1}^{n} I\{X_i \leqslant x\}$, i.e., $T(F_n)$, is regarded as a natural estimator of the parameter $\theta = T(F)$. This functional $T(\cdot)$ is defined on a space $\mathfrak{F}$ of distribution functions (d.f.) F (on R). Using a form of the Taylor expansion involving the derivative of the functional, von Mises [12] expressed $T(F_n)$ as

$$T(F_n) = T(F) + T'_F(F_n - F) + \text{Rem}(F_n - F; T), \tag{1.1}$$

where T'_F is the derivative of the functional at F and $\text{Rem}(F_n - F; T)$ is the remainder term in this first-order expansion. Note that $F_n(x)$ is based on

Received March 19, 1993; revised April 1994.

AMS 1980 subject classifications: 62E20, 62G30.

Key words and phrases: asymptotic normality, grade correlation coefficient, Hadamard differentiability, influence curve, Spearman's rank correlation coefficient, statistical functional, p-dimensional empirical process, von Mises method.

0047-259X/95 $12.00

Copyright © 1995 by Academic Press, Inc.
All rights of reproduction in any form reserved.

n independent and identically distributed random variables (i.i.d.r.v.), and that T'_F is a linear functional. Hence, $T'_F(F_n - F)$ is an average of n i.i.d.r.v.'s. For drawing statistical conclusions (in a large sample), T'_F plays the basic role, and in this context, it remains to show that $\mathrm{Rem}(F_n - F; T)$ is asymptotically negligible to the desired extent. Appropriate differentiability conditions are usually incorporated towards this verification. Since the Fréchet differentiability condition is generally too stringent, less restrictive concepts such as Gateaux̂ and Hadamard (or compact) derivatives have been considered by various people in different studies (Viz., Kallianpur [5]; Reeds [7], Fernholz [2], Gill [3], Ren and Sen [8], Ren [9], among others). In a majority of statistical applications, Hadamard differentiability usually suffices. Here we shall focus on the use of Hadamard differentiability conditions.

It is known (viz., Fernholz [2]) that a statistical functional $T(\cdot)$ induces a functional on the space $D[0, 1]$ (of right continuous functions having left-hand limits) as

$$\tau(G) = T(G \circ F), \qquad G \in D[0, 1].$$

Thus, (1.1) can be written equivalently as

$$\tau(U_n) = \tau(U) + \tau'_U(U_n - U) + \mathrm{Rem}(U_n - U; \tau), \tag{1.2}$$

where U_n is the empirical d.f. of $F(X_i)$, $1 \leqslant i \leqslant n$, and U is the classical uniform d.f. on $[0, 1]$ (i.e., $U(t) = t$, $0 \leqslant t \leqslant 1$). Using Hadamard differentiability condition (along with some other regularity conditions) Reeds [7] has shown that

$$\sqrt{n}\, \mathrm{Rem}(U_n - U; \tau) \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty \tag{1.3}$$

so that noting that $\tau'_U(U_n - U) = n^{-1} \sum_{i=1}^{n} IC(X_i; F, T)$, where $IC(x; F, T)$ is the *influence curve* of T at F, and assuming that $0 < \sigma^2 = \mathrm{Var}\{IC(X_i; F, T)\} < \infty$, one obtains

$$\sqrt{n}\, [T(F_n) - \theta] \overset{P}{\sim} \sqrt{n}\, \tau'_U(U_n - U) \xrightarrow{D} N(0, \sigma^2), \qquad \text{as} \quad n \to \infty. \tag{1.4}$$

In the multivariate case, a parameter θ $(= T(F))$ is often regarded as a functional $T(\cdot)$ on aspace $\mathfrak{F}_p$ of p-dimensional d.f.'s F on R^p, where $p \geqslant 1$. For instance, the covariance of two r.v.'s X and Y is given by

$$\sigma_{xy} = T(F) = \int\int xy\, dF(x, y) - \int\int F(x, \infty)\, F(\infty, y)\, dx\, dy,$$

and the *grade correlation coefficient* is given by

$$\rho_g = T(F) = 12 \int\int [F(x, \infty) - \tfrac{1}{2}]\, [F(\infty, y) - \tfrac{1}{2}]\, dF(x, y), \tag{1.5}$$

where F is the joint d.f. of X and Y. Using an idea similar to the one for univariate observation described above, we have $T(\tilde{F}_n)$ as the natural estimator of θ and have a form of the Taylor expansion (1.1) for $\tilde{F}_n$, where

$$\tilde{F}_n(\mathbf{x}) = n^{-1} \sum_{i=1}^{n} I\{X_{i1} \leqslant x_1, ..., X_{ip} \leqslant x_p\} \tag{1.6}$$

for $\mathbf{x} = (x_1, ..., x_p) \in R^p$, is based on n i.i.d.r.v.'s: $(X_{i1}, ..., X_{ip})$, $i = 1, ..., n$, each having the d.f. $F(\mathbf{x}) = P\{X_{i1} \leqslant x_1, ..., X_{ip} \leqslant x_p\}$.

We observe that these statistical functionals $T(\tilde{F}_n)$ induce some functionals on the space $D[0, 1]^p$ (defined in Section 2) as

$$\tau(G)(\mathbf{t}) = T(G(F_1(t_1), ..., F_p(t_p))), \qquad G \in D[0, 1]^p, \tag{1.7}$$

where $\mathbf{t} = (t_1, ..., t_p) \in [0, 1]^p$ and F_j are the marginal distributions of X_{ij}, $j = 1, ..., p$. Thus for a continuous F whith strictly increasing marginal distributions F_j, $j = 1, ..., p$, (1.1) can be written equivalently as

$$\tau(W_n) = \tau(W) + \tau'_W(W_n - W) + \text{Rem}(W_n - W; \tau), \tag{1.8}$$

where W_n is the p-dimensional empirical d.f. of sample $(U_{i1}, ..., U_{ip})$, $i = 1, ..., n$, given by $W_n(\mathbf{t}) = n^{-1} \sum_{i=1}^{n} I\{U_{i1} \leqslant t_1, ..., U_{ip} \leqslant t_p\}$ for $U_{ij} = F_j(X_{ij})$, $j = 1, ..., p$; $i = 1, ..., n$; and $\mathbf{t} = (t_1, ..., t_p) \in [0, 1]^p$, and W is a d.f. given by $W(\mathbf{t}) = P\{U_{i1} \leqslant t_1, ..., U_{ip} \leqslant t_p\}$. To draw statistical conclusions (in a large sample), we need to extend the result (1.3) to (1.8), i.e., to establish

$$\sqrt{n}\, \text{Rem}(W_n - W; \tau) \xrightarrow{P} 0, \qquad \text{as} \quad n \to \infty. \tag{1.9}$$

In Section 2, we introduce the space $D[0, 1]^p$, some notations, and the notion of Hadamard differentiability. In Section 3, we establish (1.9) and show that an estimator given by $\tau(W_n)$ is asymptotically normal if τ is Hadamard differentiable. To illustrate the use of the von Mises method, we show, in Section 4, that Spearman's rank correlation coefficient is stochastically equivalent to an estimator which is Hadamard differentiable, so that by our result (1.9), Spearman's rank correlation coefficient is asymptotically normal.

2. Preliminaries

Let $E_p = [0, 1]^p$, $\mathfrak{P} = \{\rho = (\rho_1, ..., \rho_p);\ \rho_j = 0 \text{ or } 1 \text{ for } j = 1, ..., p\}$, and define *quadrants* for every $\mathbf{t} = (t_1, ..., t_p) \in E_p$ and $\rho \in \mathfrak{P}$ by $\mathfrak{Q}(\rho, \mathbf{t}) = I(\rho_1, t_1) \times \cdots \times I(\rho_p, t_p)$, where $I(\rho_j, t_j) = [0, t_j)$, if $\rho_j = 0$; $(t_j, 1]$, if $\rho_j = 1$.

For a function f defined on E_p and $\mathbf{t} \in E_p$, $\rho \in \mathfrak{P}$, if $\lim_{n\to\infty} f(\mathbf{t}_n)$ exists with (the necessarily unique) limit $f(\mathbf{t}+0_\rho)$ for every sequence $\{\mathbf{t}_n\} \subset \mathfrak{Q}(\rho, \mathbf{t}) \neq \emptyset$ with $\mathbf{t}_n \to \mathbf{t}$, then $f(\mathbf{t}+0_\rho)$ is called the *ρ-limit* or the *quadrant limit* of f at $\mathbf{t}$. The space $D[0, 1]^p$ is the space of all real-valued functions $f: E_p \to R$ for which the ρ-limit (or the quadrant limit) of f at $\mathbf{t}$ exists for every $\mathbf{t} \in E_p$ and $\rho \in \mathfrak{P}$ with $\mathfrak{Q}(\rho, \mathbf{t}) \neq \emptyset$, and which are "continuous from above" in the sense that $f(\mathbf{t}) = f(\mathbf{t}+0_\mathbf{1})$ for $\mathbf{1} = (1, ..., 1)$.

The space $C[0, 1]^p$ is the space of all continuous functions $f: E_p \to R$. It is clear that $C[0, 1]^p$ is a subspace of $D[0, 1]^p$. More detailed discussions on the space $C[0, 1]^p$ and $D[0, 1]^p$ can be found in Neuhaus [6] or Sen [10, Chap. 2].

Let $\|\cdot\|$ denote the uniform norm. In this paper, we consider the space $(D[0, 1]^p, \|\cdot\|, \mathfrak{D})$ and the space $(C[0, 1]^p, \|\cdot\|, \mathfrak{C})$, where $\mathfrak{D}$ is the σ-field of subsets of $D[0, 1]^p$ generated by the open balls and $\mathfrak{C}$ is the σ-field of Borel subsets of $C[0, 1]^p$. Note that $(D[0, 1]^p, \|\cdot\|)$ is a (non-separable) Banach space and that $(C[0, 1]^p, \|\cdot\|)$ is a separable Banach space.

The p-dimensional empirical process considered in this paper is given by

$$W_n(\mathbf{t}) = n^{-1} \sum_{i=1}^{n} I\{U_{i1} \leqslant t_1, ..., U_{ip} \leqslant t_p\}, \tag{2.1}$$

where $\mathbf{t} = (t_1, ..., t_p) \in E_p$ and $\mathbf{U}_i = (U_{i1}, ..., U_{ip})$, $i = 1, ..., n$, are i.i.d.r.v.'s with a joint distribution function

$$W(\mathbf{t}) = P\{U_{i1} \leqslant t_1, ..., U_{ip} \leqslant t_p\}. \tag{2.2}$$

The random variables U_{ij}, $j = 1, ..., p$; $i = 1, ..., n$, are uniform $(0, 1)$ r.v.'s. Note that for $\mathbf{U}_i$, $i = 1, ..., n$, W_n is an element of $D[0, 1]^p$, and that the choice of the σ-field $\mathfrak{D}$ given above is to ensure that W_n is a random element in space $D[0, 1]^p$.

We introduce the notion of Hadamard differentiability as follows: Let V and W be the topological vector spaces and $L_1(V, W)$ be the set of continuous linear transformations from V to W. Let $\mathscr{A}$ be an open set of V.

DEFINITION. A functional $\tau: \mathscr{A} \to W$ is *Hadamard differentiable* (or *compact differentiable*) at $F \in \mathscr{A}$ if there exists $\tau'_F \in L_1(V, W)$ such that for any compact set K of V,

$$\lim_{t\to 0} \frac{\tau(F+tH) - \tau(F) - \tau'_F(tH)}{t} = 0 \tag{2.3}$$

uniformly for any $H \in K$. The linear function τ'_F is called the *Hadamard derivative* of τ at F.

For the sake of convenience, in (2.3) we usually denote the remainder term of the first-order expansion as

$$\text{Rem}(tH;\tau)=\tau(F+tH)-\tau(F)-\tau'_F(tH). \tag{2.4}$$

Particularly, for our study in this paper, we consider the functional τ defined on the space $D[0,1]^p$ and consider

$$\text{Rem}(tH;\tau)=\tau(W+tH)-\tau(W)-\tau'_W(tH) \tag{2.5}$$

for $H\in D[0,1]^p$.

3. Hadamard Differentiability Theory on $D[0,1]^p$

We establish (1.9) of Section 1 in the following theorem.

THEOREM 3.1. *Suppose τ: $D[0,1]^p\to R$ is a functional and is Hadamard differentiable at W. Then,*

$$\sqrt{n}\,\text{Rem}(W_n-W;\tau)\xrightarrow{P}0,\qquad as\quad n\to\infty. \tag{3.1}$$

Therefore,

$$\sqrt{n}\,[\tau(W_n)-\tau(W)]=\sqrt{n}\,\tau'_W(W_n-W)+o_p(1),\qquad as\quad n\to\infty. \tag{3.2}$$

The proof of Theorem 3.1 is given at the end of this section. Based on Theorem 3.1, we can easily show in the next theorem that an estimator $\hat{\theta}_n$ of θ is asymptotically normal if it can be expressed as a Hadamard differentiable functional defined on $D[0,1]^p$.

As below, we denote an influence curve of functional τ defined on $D[0,1]^p$ at W as

$$IC_p(\mathbf{x};W,\tau)=\frac{d}{dt}\tau(W+t(\delta_{\mathbf{x}}-W))|_{t=0},$$

where $\delta_{\mathbf{x}}$ is the d.f. of the point mass one at $\mathbf{x}\in E_p$.

THEOREM 3.2. *Let $\hat{\theta}_n=\tau(W_n)$ be an estimator of $\theta=\tau(W)$. If the functional τ: $D[0,1]^p\to R$ is Hamadamard differentiable at W and if $0<\sigma^2=\text{Var}\{IC_p(\mathbf{U}_i;W,\tau)\}<\infty$, then*

$$\sqrt{n}\,[\hat{\theta}_n-\theta]\xrightarrow{D}N(0,\sigma^2),\qquad as\quad n\to\infty$$

when $IC_p(\mathbf{x};W,\tau)$ is centered so that $\int IC_p(\mathbf{x};W,\tau)\,dW=0$.

Proof. By Theorem 3.1, we have that as $n \to \infty$,

$$\sqrt{n}\,[\hat{\theta}_n - \theta] = \sqrt{n}\,\tau'_W(W_n - W) + o_p(1) = n^{-1/2} \sum_{i=1}^{n} \tau'_W(\delta_{\mathbf{U}_i} - W) + o_p(1)$$

because τ'_W is a linear function. Since by the definition of the Hadamard derivative,

$$IC_p(\mathbf{x};\, W, \tau) = \lim_{t \to 0} \frac{\tau(W + t(\delta_{\mathbf{x}} - W)) - \tau(W)}{t} = \tau'_W(\delta_{\mathbf{x}} - W),$$

the proof follows from the central limit theorem, Slutsky's lemma, and the fact:

$$\sqrt{n}\,[\hat{\theta}_n - \theta] = n^{-1/2} \sum_{i=1}^{n} IC_p(\mathbf{U}_i;\, W, \tau) + o_p(1), \qquad \text{as} \quad n \to \infty. \quad \blacksquare$$

Before proving Theorem 3.1, we first establish the following lemmas with the proofs given in the Appendix. Since the space $D[0, 1]^p$ is not separable, we will consider a continuous version of W_n, say $W_n^* \in C[0, 1]^p$, so that the tightness of the sequence of probability measures corresponding to W_n^* can be obtained later. We note that the rectangles

$$R_i = \{\mathbf{t} \in E_p \mid U_{i1} \leqslant t_1, ..., U_{ip} \leqslant t_p\}, \qquad i = 1, ..., n,$$

divide E_p into finite pieces and that $W_n(\mathbf{t})$ is a simple function, i.e., the value of $W_n(\mathbf{t})$ is a constant in each piece divided by R_i's. Hence, the continuous version of W_n, denoted by W_n^*, may be obtained by smoothing W_n through some hyperplanes according to R_i's. We have the following lemma on the difference between W_n and W_n^*.

LEMMA 3.3. *With probability one, any edges of different rectangles in the interior of E_p do not overlap or partially overlap. Therefore, with probability one*

$$\| W_n - W_n^* \| \leqslant p/n. \tag{3.3}$$

We note that such continuous version W_n^* of W_n is a continuous function defined on E_p for given $\mathbf{U}_i$, $i = 1, ..., n$, and is piecewise linear. We also note that for any fixed $\mathbf{t} \in E_p$, the cross section $W_n^*(\mathbf{t})$ is a simple function of $\mathbf{U}_i$, $i = 1, ..., n$. Hence, $W_n^*(\mathbf{t})$ is a random variable. Thus, W_n^* is a random element of the separable space $C[0, 1]^p$.

LEMMA 3.4. *Let $\{P_n; n \geq 1\}$ be the sequence of probability measures corresponding to $T_n(\mathbf{t}) = \sqrt{n}\,[W_n^*(\mathbf{t}) - W(\mathbf{t})]$ for $\mathbf{t} \in E_p$. Then, $\{P_n; n \geq 1\}$ is relatively compact.*

The following lemma is a generalized version on $D[0, 1]^p$ of Lemma 4.3.1 of Fernholz [2]. The proof follows line by line of the one for Lemma 4.3.1 (Fernholz [2]). Let K be a set in $D[0, 1]^p$ and $H \in D[0, 1]^p$, we denote

$$\operatorname{dist}(H, K) = \inf_{G \in K} \|H - G\|.$$

LEMMA 3.5. *Let Q: $D[0, 1]^p \times R \to R$ and suppose that for any compact set K in $D[0, 1]^p$.*

$$\lim_{t \to 0} Q(H, t) = 0 \tag{3.4}$$

uniformly for $H \in K$. Let $\varepsilon > 0$, and let δ_n be a sequence of numbers such that $\delta_n \to 0$. Then, for any compact set K in $D[0, 1]^p$ there exists a positive integer N such that for any $n > N$, $\operatorname{dist}(H, K) \leq p\delta_n$ implies that

$$|Q(H, \delta_n)| < \varepsilon.$$

Proof of Theorem 3.1. By Lemma 3.4, $\{P_n; n \geq 1\}$ is relatively compact in $C[0, 1]^p$, where $P_n\{A\} = P\{T_n \in A\}$. Since $C[0, 1]^p$ is complete and separable, by Prohorov's theorem (Billingsley [1, Theorem 6.2]), $\{P_n; n \geq 1\}$ is tight, i.e., for any $\varepsilon > 0$, there exists a compact set K in $C[0, 1]^p$ such that for all $n \geq 1$,

$$P\{T_n \in K\} > 1 - \varepsilon. \tag{3.5}$$

Since $C[0, 1]^p$ is a subspace of $D[0, 1]^p$, K is also a compact set in $D[0, 1]^p$. If $T_n = \sqrt{n}\,[W_n^* - W] \in K$ and $\|W_n - W_n^*\| \leq p/n$, then

$$\operatorname{dist}(\sqrt{n}\,[W_n - W], K) \leq p/\sqrt{n}.$$

Hence, by (3.3) and (3.5), we have

$$P\{\operatorname{dist}(\sqrt{n}\,[W_n - W], K) \leq p/\sqrt{n}\} > 1 - \varepsilon.$$

Since τ: $D[0, 1]^p \to R$ is Hadamard differentiable at W, by the definition of Hadamard differentiability, (3.4) holds for $Q(H, t) = \operatorname{Rem}(tH; \tau)/t$. By Lemma 3.5, there exists N such that for $n > N$, $\operatorname{dist}(H, K) \leq p/\sqrt{n}$ implies

$$|\sqrt{n}\operatorname{Rem}(H/\sqrt{n}; \tau)| < \varepsilon.$$

Therefore for $n > N$ and $H = \sqrt{n}\,[W_n - W]$, we have

$$P\{\sqrt{n}\,\mathrm{Rem}(W_n - W; \tau) < \varepsilon\} > 1 - \varepsilon. \quad \blacksquare$$

4. Asymptotic Normality of Spearman's Rank Correlation Coefficient

Let $(X_1, Y_1), ..., (X_n, Y_n)$ be a random sample from $F(x, y)$ which is continuous with strictly increasing and continuous marginal d.f. F_x and F_y. Let $R_1, ..., R_n$ and $S_1, ..., S_n$ denote the ranks of the X_i's and Y_i's, respectively. A popular estimate of the strength of the association between the two characteristics in the population from which the sample is drawn is given by

$$r_s = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2 \sum_{i=1}^n (S_i - \bar{S})^2}}, \tag{4.1}$$

where $\bar{R} = n^{-1} \sum_{i=1}^n R_i$ and $\bar{S} = n^{-1} \sum_{i=1}^n S_i$. Known as *Spearman's rank correlation coefficient*, r_s was proposed by Spearman [11]. We will use our results on Hadamard differentiability given in Section 3 to prove the asymptotic normality of r_s.

First, we observe that by virtue of the assumed continuity of the marginal d.f.'s, ties among X_i's or Y_i's are neglected in probability, and hence $\bar{R} = \bar{S} = (n+1)/2$ and $\sum_{i=1}^n (R_i - \bar{R})^2 = \sum_{i=1}^n (S_i - \bar{S})^2 = n(n^2 - 1)/12$. Therefore we have

$$\begin{aligned} r_s &= \frac{12}{n(n^2-1)} \sum_{i=1}^n \left(R_i - \frac{n+1}{2}\right)\left(S_i - \frac{n+1}{2}\right) \\ &= \frac{12(n+1)}{n(n-1)} \sum_{i=1}^n \left(\frac{R_i}{n+1} - \frac{1}{2}\right)\left(\frac{S_i}{n+1} - \frac{1}{2}\right). \end{aligned}$$

If we denote

$$F_n(x, y) = n^{-1} \sum_{i=1}^n I\{X_i \leqslant x,\ Y_i \leqslant y\}$$

$$G_n(x) = n^{-1} \sum_{i=1}^n I\{X_i \leqslant x\} = F_n(x, \infty)$$

$$H_n(y) = n^{-1} \sum_{i=1}^n I\{Y_i \leqslant y\} = F_n(\infty, y),$$

then a natural estimator of the grade correlation coefficient

$$\rho_g = 12 \int\int [F_x(x) - \tfrac{1}{2}][F_y(y) - \tfrac{1}{2}]\, dF(x, y)$$

is given by

$$\begin{aligned} r_g &= 12 \int\int [G_n(x) - \tfrac{1}{2}][H_n(y) - \tfrac{1}{2}]\, dF_n(x, y) \\ &= 12n^{-1} \sum_{i=1}^{n} (G_n(X_i) - \tfrac{1}{2})(H_n(Y_i) - \tfrac{1}{2}). \end{aligned} \tag{4.2}$$

It is easy to verify that

$$r_s = \frac{(n+1)}{(n-1)} r_g - \frac{12(2n+1)}{n^3(n^2-1)} \sum_{i=1}^{n} R_i S_i + \frac{6(n+1)}{n(n-1)}. \tag{4.3}$$

We note that if, in (2.1), we let $U_i = U_{i1} = F_x(X_i)$ and $V_i = U_{i2} = F_y(Y_i)$, then for $(s, t) \in E_2$,

$$W_n(s, t) = F_n(F_x^{-1}(s), F_y^{-1}(t)),$$

$$W(s, t) = F(F_x^{-1}(s), F_y^{-1}(t)),$$

and the estimator r_g of ρ_g can be expressed as a functional of W_n:

$$r_g = \Psi(W_n) = 12 \int_0^1 \int_0^1 [W_n(s, 1) - \tfrac{1}{2}][W_n(1, t) - \tfrac{1}{2}]\, dW_n(s, t). \tag{4.4}$$

Note that $W(s, 1) = s$ and $W(1, t) = t$; hence the grade correlation coefficient ρ_g can be expressed as the same functional of W:

$$\rho_g = \Psi(W) = 12 \int_0^1 \int_0^1 [s - \tfrac{1}{2}][t - \tfrac{1}{2}]\, dW(s, t). \tag{4.5}$$

Next, we show that this functional Ψ is Hadamard differentiable at W. Let

$$E = \left\{ H \in D[0, 1]^2; \int_0^1 \int_0^1 |dH| \leqslant C \right\}$$

for some constant $C \geqslant 1$ (see Gill [3, p. 109] for the discussions on E).

THEOREM 4.1. *Let Ψ: $D[0,1]^2 \to R$ be a functional given by*

$$\Psi(H) = 12 \int_0^1 \int_0^1 [H(s,1) - \tfrac{1}{2}][H(1,t) - \tfrac{1}{2}]\, dH(s,t),$$

where $\Psi(H)$ is defined for $H \in E$. Then, Ψ is a Hadamard differentiable at W with Hadamard derivative

$$\Psi'_W(H) = 12 \left(\int_0^1 \int_0^1 (s - \tfrac{1}{2})(t - \tfrac{1}{2})\, dH(s,t) + \int_0^1 \int_0^1 [(s - \tfrac{1}{2})\, H(1,t) \right.$$

$$\left. + (t - \tfrac{1}{2})\, H(s,1)]\, dW(s,t) \right).$$

Proof. By (4) of Gill, it suffices to show that

$$\frac{\operatorname{Rem}(t_n H_n; \Psi)}{t_n} = \frac{\Psi(W + t_n H_n) - \Psi(W) - \Psi'_W(t_n H_n)}{t_n} \to 0$$

for every sequence $t_n \in R^+$, $H_n \in D[0,1]^2$ with $t_n \to 0$, $H_n \to H$, as $n \to \infty$, and $H \in D[0,1]^2$, $G_n = W + t_n H_n \in E$. Since $W(s,1) = s$ and $W(1,t) = t$, we have

$$\frac{\operatorname{Rem}(t_n H_n; \Psi)}{t_n} = 12 t_n \int_0^1 \int_0^1 H_n(s,1)\, H_n(1,t)\, dW(s,t)$$

$$+ 12 \int_0^1 \int_0^1 \{\psi_{H_n}(s,t) + t_n H_n(s,1)\, H_n(1,t)\}\, d[G_n - W],$$

where $\psi_{H_n}(s,t) = [s - \tfrac{1}{2}]\, H_n(1,t) + [t - \tfrac{1}{2}]\, H_n(s,1)$. Note that

$$\left| t_n \int_0^1 \int_0^1 H_n(s,1)\, H_n(1,t)\, dW(s,t) \right| \leqslant M t_n$$

and

$$\left| \int_0^1 \int_0^1 t_n H_n(s,1)\, H_n(1,t)\, d[G_n - W] \right|$$

$$\leqslant M t_n \left\{ \int_0^1 \int_0^1 |dG_n| + \int_0^1 \int_0^1 |dW| \right\} \leqslant 2MC t_n,$$

for some constant $M > 0$. Also, note that

$$\left|\int_0^1 \int_0^1 \psi_{H_n}(s, t)\, d[G_n - W]\right|$$
$$\leqslant \|\psi_{H_n} - \psi_H\| \left\{\int_0^1 \int_0^1 |dG_n| + \int_0^1 \int_0^1 |dW|\right\}$$
$$+\left|\int_0^1 \int_0^1 \psi_H(s, t)\, d[G_n - W]\right|,$$
$$\leqslant 2C\, \|\psi_{H_n} - \psi_H\| + \left|\int_0^1 \int_0^1 \psi_H(s, t)\, d[G_n - W]\right|;$$

hence from $\|\psi_{H_n} - \psi_H\| \to 0$, as $n \to \infty$, it suffices to show that

$$\left|\int_0^1 \int_0^1 \psi_H(s, t)\, d[G_n - W]\right| \to 0, \qquad \text{as} \quad n \to \infty.$$

Since $\psi_H \in D[0, 1]^2$, from Lemma 1.5 of Neuhaus [6], we have that for any $\varepsilon > 0$, there exists a simple function $\psi^{\varepsilon}_H(\cdot) = \sum_{k=1}^{N_\varepsilon} \alpha_k I_{R_k}(\cdot)$, where N_ε is a positive integer, R_k are rectangles, $I_{R_k}(\cdot)$ is the indicator of R_k, and $\alpha_k \in R$, such that

$$\|\psi_H - \psi^{\varepsilon}_H\| \leqslant \varepsilon.$$

The proof follows from

$$\left|\int_0^1 \int_0^1 \psi_H d[G_n - W]\right| \leqslant \|\psi_H - \psi^{\varepsilon}_H\| \left\{\int_0^1 \int_0^1 |dG_n| + \int_0^1 \int_0^1 |dW|\right\}$$
$$+\left|\int_0^1 \int_0^1 \psi^{\varepsilon}_H d[G_n - W]\right|$$
$$\leqslant 2\varepsilon C + \left|\sum_{k=1}^{N_\varepsilon} \alpha_k \left\{\int\int_{R_k} dG_n - \int\int_{R_k} dW\right\}\right|$$
$$\leqslant 2\varepsilon C + 4\, \|\psi^{\varepsilon}_H\|\, \|G_n - W\|$$
$$= 2\varepsilon C + 4\, \|\psi^{\varepsilon}_H\|\, \|H_n\|\, t_n \to 0, \qquad \text{as} \quad n \to \infty. \quad \blacksquare$$

THEOREM 4.2. *Assume $F(x, y)$ is a continuous d.f. with strictly increasing and continuous marginal d.f. F_x and F_y. Then, we have*

$$\sqrt{n}\,[r_g - \rho_g] \xrightarrow{D} N(0, \sigma^2_s), \qquad \text{as} \quad n \to \infty, \tag{4.6}$$

where $\sigma_s^2 = \mathrm{Var}\{Z_i\}$ *for*

$$Z_i = 12\left\{[F_x(X_i)-\tfrac{1}{2}][F_y(Y_i)-\tfrac{1}{2}] + \int\int \{[F_x(x)-\tfrac{1}{2}]\, I\{Y_i \leqslant y\} + [F_y(y)-\tfrac{1}{2}]\, I\{X_i \leqslant x\}\}\, dF(x,y)\right\}.$$

Therefore, we have

$$\sqrt{n}\,[r_s - \rho_g] \xrightarrow{D} N(0, \sigma_s^2), \qquad as \quad n \to \infty. \tag{4.7}$$

Proof. We first note that it suffices to establish (4.6), because by (4.3) we have

$$\begin{aligned}\sqrt{n}\,(r_s - r_g) &= \sqrt{n}\left(\frac{n+1}{n-1} - 1\right) r_g - \frac{12(2n+1)\sqrt{n}}{n^3(n^2-1)} \sum_{i=1}^{n} R_i S_i + \frac{6(n+1)}{\sqrt{n}(n-1)} \\ &= \frac{2}{n-1}\sqrt{n}\,[r_g - \rho_g] + \frac{2\sqrt{n}}{n-1}\rho_g - \frac{12(2n+1)\sqrt{n}}{n^3(n^2-1)} \\ &\quad \times \sum_{i=1}^{n} R_i S_i + \frac{6(n+1)}{\sqrt{n}\,(n-1)},\end{aligned}$$

where

$$\begin{aligned}\frac{12(2n+1)\sqrt{n}}{n^3(n^2-1)} \sum_{i=1}^{n} R_i S_i &\leqslant \frac{12(2n+1)\sqrt{n}}{n^3(n^2-1)} \sqrt{\textstyle\sum_{i=1}^{n} R_i^2}\, \sqrt{\textstyle\sum_{i=1}^{n} S_i^2} \\ &= \frac{2(2n+1)^2\sqrt{n}}{n^2(n-1)},\end{aligned}$$

and $\sqrt{n}\,[r_s - \rho_g] = \sqrt{n}\,[r_s - r_g] + \sqrt{n}\,[r_g - \rho_g]$.

By Theorem 4.1, we know that the functional Ψ is Hadamard differentiable at W. By (4.4), (4.5), and Theorem 3.1, we have, as $n \to \infty$,

$$\sqrt{n}\,[r_g - \rho_g] = \sqrt{n}\,[\Psi(W_n) - \Psi(W)] = \sqrt{n}\,\Psi'_W(W_n - W) + o_p(1). \tag{4.8}$$

Since

$$\begin{aligned}\Psi'_W(W_n - W) &= 12\int_0^1\int_0^1 (s-\tfrac{1}{2})(t-\tfrac{1}{2})\, dW_n(s,t) \\ &\quad - 12\int_0^1\int_0^1 (s-\tfrac{1}{2})(t-\tfrac{1}{2})\, dW(s,t) \\ &\quad + 12\int_0^1\int_0^1 \{(s-\tfrac{1}{2})\, W_n(1,t) + (t-\tfrac{1}{2})\, W_n(s,1)\}\, dW(s,t)\end{aligned}$$

$$
\begin{aligned}
&-12\int_0^1\int_0^1 \{(s-\tfrac{1}{2})\,t+(t-\tfrac{1}{2})\,s\}\,dW(s,t)\\
&=12n^{-1}\sum_{i=1}^{n}(U_i-\tfrac{1}{2})(V_i-\tfrac{1}{2})-12\int_0^1\int_0^1 (s-\tfrac{1}{2})(t-\tfrac{1}{2})\,dW(s,t)\\
&\quad+12n^{-1}\sum_{i=1}^{n}\int_0^1\int_0^1 \{(s-\tfrac{1}{2})\,I\{V_i\leqslant t\}\\
&\quad+(t-\tfrac{1}{2})\,I\{U_i\leqslant s\}\}\,dW(s,t)\\
&\quad-12\int_0^1\int_0^1 \{(s-\tfrac{1}{2})\,t+(t-\tfrac{1}{2})\,s\}\,dW(s,t)\\
&=12n^{-1}\sum_{i=1}^{n}\Big\{(U_i-\tfrac{1}{2})(V_i-\tfrac{1}{2})+\int_0^1\int_0^1 \{(s-\tfrac{1}{2})\,I\{V_i\leqslant t\}\\
&\quad+(t-\tfrac{1}{2})\,I\{U_i\leqslant s\}\}\,dW(s,t)\Big\}\\
&\quad-12\int_0^1\int_0^1 \{(s-\tfrac{1}{2})(t-\tfrac{1}{2})+(s-\tfrac{1}{2})\,t+(t-\tfrac{1}{2})\,s\}\,dW(s,t),
\end{aligned}
$$

then

$$
\Psi'_W(W_n-W)=n^{-1}\sum_{i=1}^{n}[Z_i-E\{Z_i\}],
$$

where

$$
\begin{aligned}
Z_i&=12\Big\{(U_i-\tfrac{1}{2})(V_i-\tfrac{1}{2})+\int_0^1\int_0^1 \{(s-\tfrac{1}{2})\,I\{V_i\leqslant t\}\\
&\quad+(t-\tfrac{1}{2})\,I\{U_i\leqslant s\}\}\,dW(s,t)\Big\}\\
&=12\Big\{[F_x(X_i)-\tfrac{1}{2}][F_y(Y_i)-\tfrac{1}{2}]+\int\int \{[F_x(x)-\tfrac{1}{2}]\,I\{Y_i\leqslant y\}\\
&\quad+[F_y(y)-\tfrac{1}{2}]\,I\{X_i\leqslant x\}\}\,dF(x,y)\Big\}
\end{aligned}
$$

are i.i.d.r.v.'s. Therefore, by (4.8) we have that as $n\to\infty$,

$$
\sqrt{n}\,[r_g-\rho_g]=n^{-1/2}\sum_{i=1}^{n}[Z_i-E\{Z_i\}]+o_p(1)\xrightarrow{D}N(0,\sigma_s^2),
$$

where $\sigma_s^2=\operatorname{Var}\{Z_i\}$. ∎

Remark. Hoeffding [4] has an elegant proof of the asymptotoc normality of r_s through explicit representation in terms of two U-statistics. Although in general U-statistics are not Hadamard differentiable, our alternative formulation does provide a simple proof of the asymptotic normality for the U-statistics with bounded kernel. Additionally, if instead of r_s, we consider a general score statistic $\int J(G_n(x))\, L(H_n(y))\, dF_n(x, y)$, where the score functions $J(\cdot)$ and $L(\cdot)$ are of bounded variation, we may not have an explicit representation in terms of U-statistics, but our Hadamard differentiability condition holds. This shows the adaptability of the proposed method in a wider class of statistics.

APPENDIX

Proof of Lemma 3.3. For the sake of simplicity, we consider the case of $p = 2$. Note that the largest jump of W_n is no larger than $2/n$, if any edges of different rectangles in the interior of E_2 do not overlap or partially overlap. The proof follows from the fact that with probability one, $U_{i1} \neq U_{j1}$ and $U_{i2} \neq U_{j2}$ for $i \neq j$, because U_{ij} are continuous random variables. ∎

Proof of Lemma 3.4. Note that $T_n(\mathbf{t}) \in C[0, 1]^p$, $W_n(\mathbf{0}) = 0$, and $W(\mathbf{0}) = 0$. Hence, by Lemma 3.3, we have

$$|T_n(\mathbf{0})| = \sqrt{n}\, |W_n^*(\mathbf{0})| \leqslant p/\sqrt{n},$$

with probability one. Therefore, $P_n^0 = P\pi_0^{-1}$ converges in distribution. By virtue of Neuhaus [6, discussion on p. 1290–1291], it suffices to show that for any $\varepsilon > 0$,

$$\lim_{\delta \to 0} \overline{\lim_{n \to \infty}}\, P\{\omega_{T_n}(\delta) \geqslant \varepsilon\} = 0, \tag{A1}$$

where $\omega_{T_n}(\delta) = \sup\{|T_n(\mathbf{s}) - T_n(\mathbf{t})|;\ |\mathbf{s} - \mathbf{t}| \leqslant \delta\}$.

Since for any $\mathbf{s}, \mathbf{t} \in E_p$, we have

$$\begin{aligned} |T_n(\mathbf{s}) - T_n(\mathbf{t})| \leqslant{}& \sqrt{n}\, |[W_n^*(\mathbf{s}) - W_n^*(\mathbf{t})] - [W_n(\mathbf{s}) - W_n(\mathbf{t})]| \\ &+ \sqrt{n}\, |[W_n(\mathbf{s}) - W_n(\mathbf{t})] - [W(\mathbf{s}) - W(\mathbf{t})]|, \end{aligned}$$

then by (3.3), we have that for any $\delta > 0$ and $T_n' = \sqrt{n}\, [W_n - W]$,

$$\omega_{T_n}(\delta) \leqslant 2p/\sqrt{n} + \omega_{T_n'}(\delta)$$

with probability one. Therefore (A1) follows from

$$\lim_{\delta \to 0} \overline{\lim_{n \to \infty}} \ \omega_{T_n'}(\delta) = 0$$

which, in turn, follows from Neuhaus [6, discussion on p. 1292–1294]. ∎

ACKNOWLEDGMENTS

The authors are grateful to the referee for his/her useful comments on the earlier draft of the manuscript.

REFERENCES

[1] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
[2] FERNHOLZ, L. T. (1983). *Von Mises Calculus for Statistical Functional*. Lecture Notes in Statistics, Vol. 19. Springer-Verlag, New York.
[3] GILL, R. D. (1989). Non- and semi-parametric maximum likelihood estimators and the von Mises method—Part 1. *Scand. J. Statist.* **16** 97–128.
[4] HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.
[5] KALLIANPUR, G. (1963). Von Mises function and maximum likelihood estimation. *Sankhyā A* **23** 149–158.
[6] NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285 1295.
[7] REEDS, J. A. (1976). *On the Definition of von Mises Functional*. Ph.D. dissertation, Harvard University, Cambridge, MA.
[8] REN, J., AND SEN, P. K. (1991). On Hadamard differentiability of extended statistical functional. *J. Multivariate Anal.* **39** 30–43.
[9] REN, J. (1994). On Hadamard differentiability and its application to *R*-estimation in linear models. *Statist. Decisions* **12**, No. 1, 1–22.
[10] SEN, P. K. (1981). *Sequential Nonparametrics*. Wiley, New York.
[11] SPEARMAN, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* **15** 72–101.
[12] VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309–348.